High School Choice and Student Sorting by Ability

Phil Leonard, McMaster University
Public Lecture to:
Johnson-Shoyama Graduate School of Public Policy
February 9, 2012

Outline of Presentation

- Introduction – What are the issues?
- Literature and where my paper fits
- Theoretical Model
- The datasets
- Methodology
- Descriptive statistics
- Regression results
- Summary

Introduction – School Choice

- Basic concept: if students can choose from more schools, their outcomes will be improved
  - With more choice, schools may compete for students
    - Students will leave under-performing schools to go to higher-performing schools
  - Possible better matching between students and schools
  - Major papers: Hoxby (2000), Rothstein (2005), Belfield and Levin (2002)
  - Canadian data: Card, Dooley and Payne (2010), Leonard (2010)

Two types of choice

- Tiebout choice
  - choice of residential location to provide access to preferred schools / school boards (e.g. Hoxby 2000)
- Choice of school / school board given residential location
  - Policies such as vouchers or “open enrolment” intended to increase this type of choice
  - My paper focusses on this second type of choice

Sorting into high schools

- Literature is largely based on US context
  - Moving low-income students from inner-city (low-performing) schools to higher-performing schools
  - Moving students from (low-performing) public schools to higher-performing private schools (vouchers)
- Hoxby (2000) and others – make suggestion that increased choice/ competition may result in decreasing heterogeneity of students within schools

More choice may lead to more sorting

- If best/most motivated students switch schools, weaker students may be left behind at the weaker schools
  - Peer groups may play an important role in education
  - While this may benefit the stronger students who switch, it may be detrimental to weaker students who stay
  - Cullen, Jacob and Levitt (2005) – Chicago Public Schools open enrollment program

Sorting is a concern if…

- Altonji, Huang and Taber (2011) show that for cream skimming to have a negative effect:
  - There must be variation in student ability within schools
  - Students of greater ability must be more likely to switch to private (or higher quality) schools
  - The ability of peers must influence an individual's outcomes

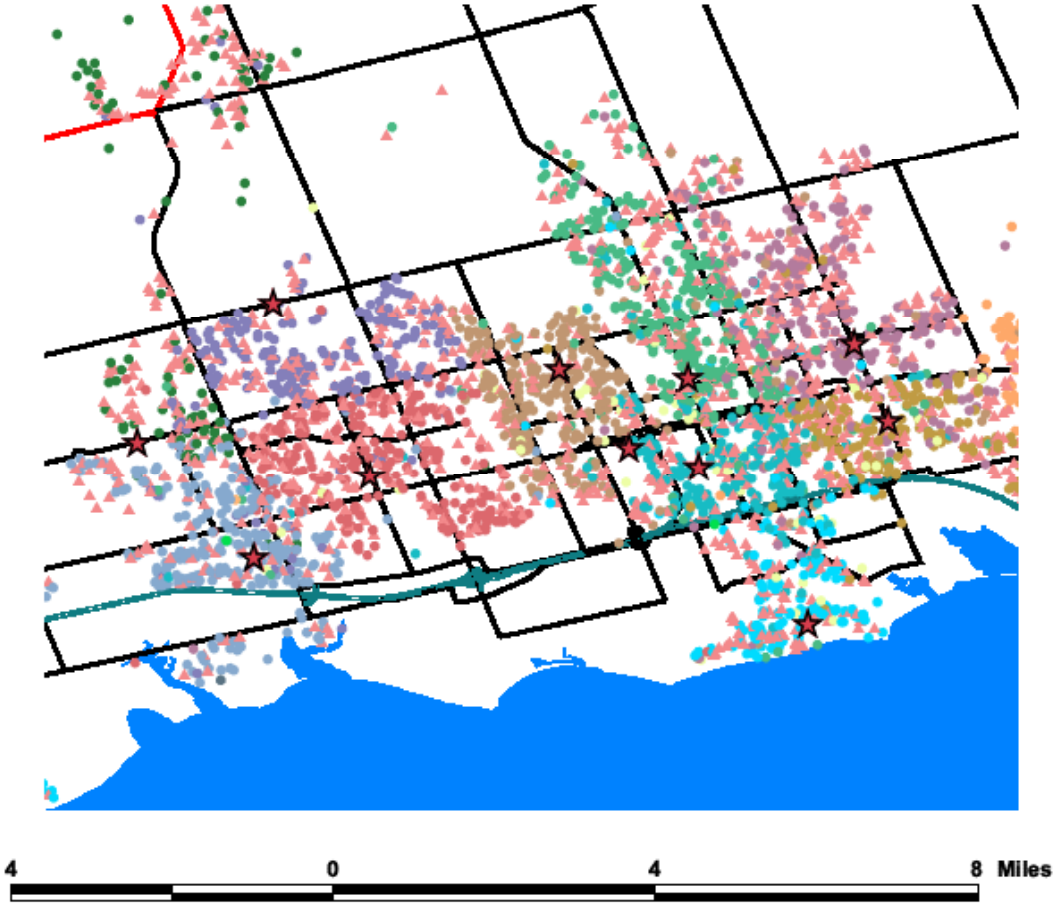
My Paper

- Goal of my paper is to examine sorting in Ontario, Canada context at time students choose their high school
  - Transition to HS likely time for sorting to occur
  - Measure choice by counting the number of schools accessible from student's residence
  - Are students of higher ability (as measured by grade 6 test scores) more likely to opt out of their assigned high schools?
  - Do students of higher ability react differently to increased choice/competition?
  - Answers first two of three criteria of Altonji, Huang, and Taber (2011)

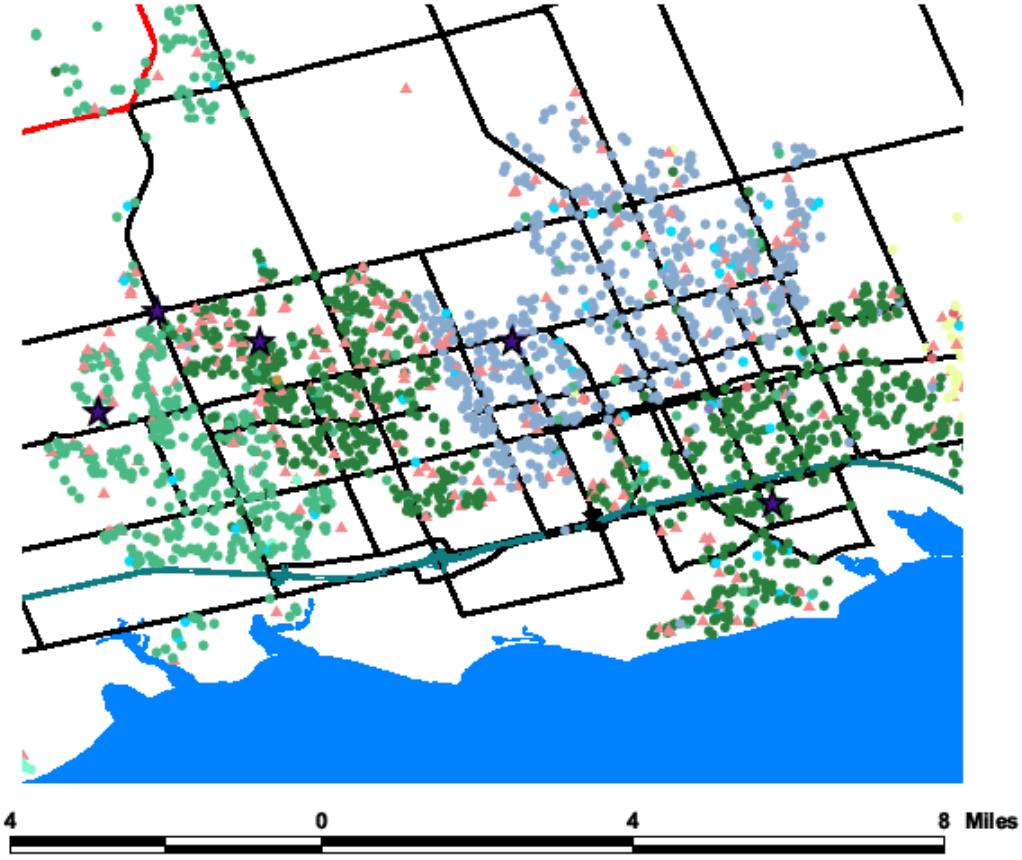
Toronto Area context

- Two (four if include French – I don't) publicly-funded school boards in each jurisdiction (Catholic and Public)
- In theory, students are assigned to a high school based on their residence
- In practice, students have a great deal of choice of which high school to attend
  - At least have choice of public vs. Catholic HS
  - Board may have explicit rules with respect to "open enrollment"
  - See students from the same neighbourhood in downtown Toronto attending 20 or more different high schools

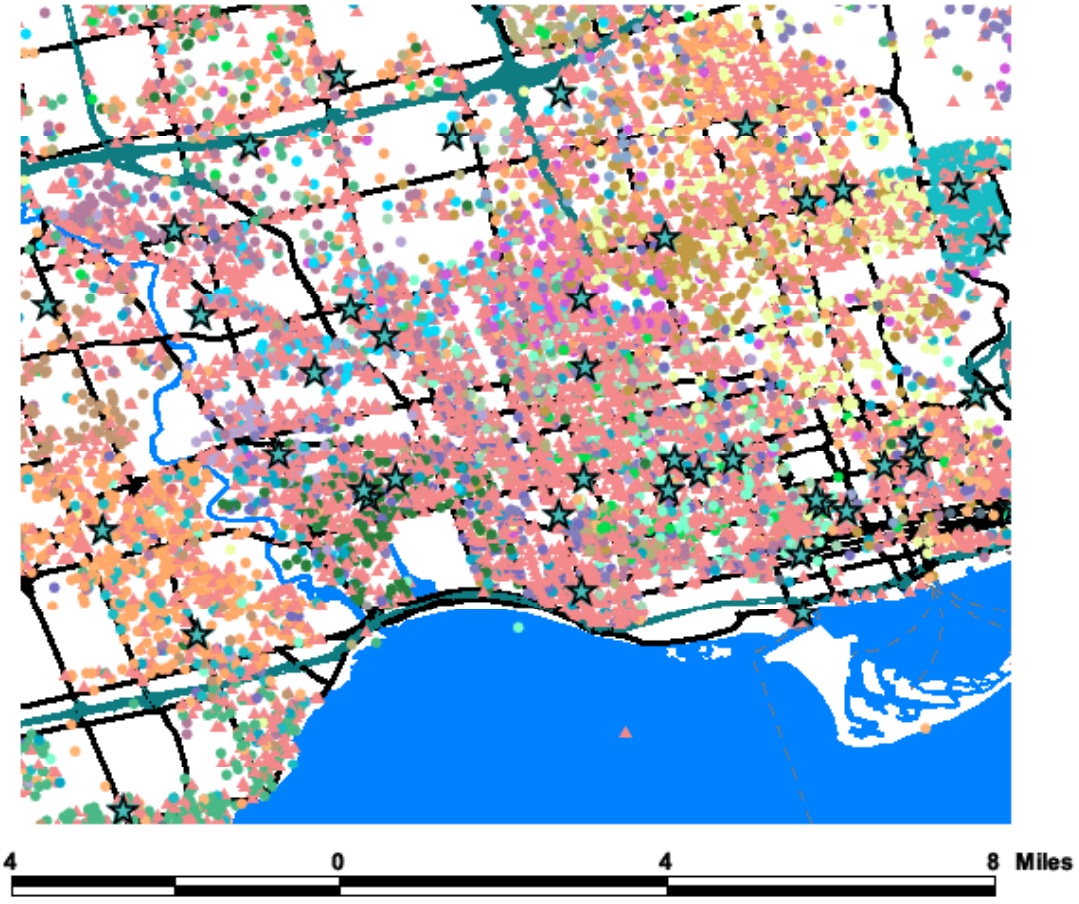
Oshawa Public Highschools



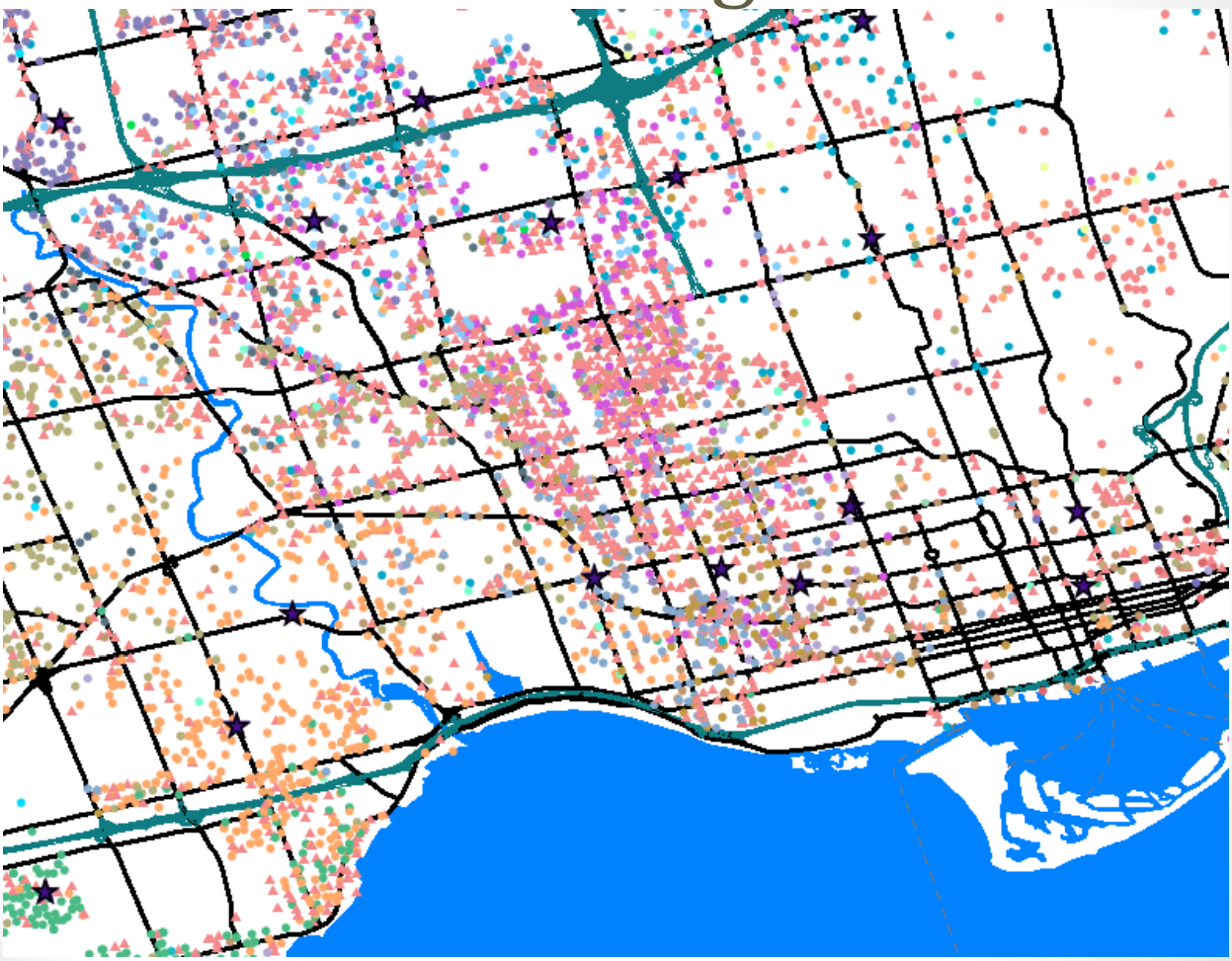
Oshawa Catholic Highschools



Toronto Public Highschools



Toronto Catholic High Schools



Theoretical Model

- Based on Chakrabarti (2009)
- Households, defined by their income and ability, (y, α) choose between their assigned high school, and a set of neighbouring alternative public schools
  - Assigned public school has quality Q_A
  - Quality of neighbouring schools: (Q_1, Q_2, … Q_N)
    - Assume that at least one of $Q_1 \ldots Q_N > Q_A$ (or nobody would ever switch)
  - School qualities determined exogenously

Household utility

- Household utility is a function of a numeraire good, x, ability, α, and school quality, θ
  - $U(x, \theta, \alpha) = h(x) + \alpha\, u(\theta)$
  - More able families (students) gain more from school quality (higher marginal valuation)
  - Increasing and concave in x and θ

School choice

- Student can attend assigned school with no cost to household
  - $U_A (x, \theta, \alpha) = h(y) + \alpha\, u (Q_A)$
- Opting to attend a neighbouring school results in direct costs, c_1, (such as travel costs, new books/uniforms) and indirect costs, c_2, such as acclimatization or travel time
  - $U_N (x, \theta, \alpha) = h(y - c_2) - c_1 + \alpha\, u (Q_N)$

Two school equilibrium

- Consider the choice between the assigned school and a single alternative school
- Household will opt out of assigned school only if difference, D, in utilities is positive where
  - $D = h(y - c_2) - c_1 + \alpha\, u(Q_N) - h(y) - \alpha\, u(Q_A)$
  - $\delta D/\delta\alpha = u(Q_N) - u(Q_A) > 0$, so students of greater ability are more likely to switch
  - $\delta D/\delta y = h'(y - c_2) - h'(y) > 0$, so higher income families/students are more likely to switch

Multiple school equilibrium

- Let neighbouring school which maximizes household utility (among neighbouring schools) have quality and costs (Q^*, C_1^*, C_2^*)
  - As N increases (more schools accessible), maximum utility can only increase as either Q^* increases or costs decrease
- As before,
  - $D = h(y - c2^*) - c1^* + \alpha u(Q^*) - h(y) - \alpha u(QA)$
  - So again, $\delta D/\delta\alpha$, $\delta D/\delta y > 0$, so families of high income or ability are more likely to switch
  - Families of higher ability also react most strongly to increases in choice (N) since $\delta D/\delta Q^* = \alpha u'(Q^*) > 0$

The main dataset

- Linked Gr. 6 (2004, 2005) to Gr. 9 EQAO (2007, 2008) database
  - Contains student test scores
    - Gr. 6 math, reading, writing (scored 0 to 4)
    - Gr. 9 math (academic or applied – each scored 0 to 4)
    - Some student-level info (gifted, ESL, spec.ed.)
  - School identifiers for both grades
    - Can link to information on school, such as: school size, board type, average score for school,
    - Can also link to census information on the forward sortation area (fsa) surrounding the school

Other Datasets

- School Attendance by postal code
  - Allows me to create attendance travel zones for each high school
  - Same school travel zones as my previous paper
- Census data
  - Use data on the fsa surrounding the elementary schools (for demographic information on parents)
  - I make assumption that students reside close to their elementary schools
- Hand checking – using school board websites, I hand checked the assigned high school for students at each of 1000 GTA elementary schools

The sample

- Focus on students in eight Toronto area school boards
  - Toronto, Durham, York and Peel (Catholic and Public Boards)
  - 99,800 student-level observations (2 cohorts in grade 6 in 2004, 2005)
  - 29,000 observations if restrict distance to 5 km of board boundary (for IV methodology)


Greater Toronto Area
Georgina
Brock
East Gwillimbury
Durham
Uxbridge
Scugog
1
2
3
Peel
York
King
Caledon
Vaughan
Markham
Pickering
Ajax
Clarington
Brampton
Halton Hills
Halton
City of Toronto
Mississauga
Milton
Oakville
Lake Ontario
N

Proposed methodology

- Would like to estimate the following:
- $Prob(S_i=1) = C_i + Gr6_i + C_i * Gr6_i + Indiv_i + Sch6_i + Census_i + e_i$
  - S_i – binary variable equal to 1 if student opted out of his/her assigned high school
  - C_i – choice/competition; count of high schools accessible to grade 6 student i
  - $Gr6_i$ – grade 6 math / reading score of student i
  - $Indiv_i$ – vector of characteristics of student i (gender, ESL status, gifted, spec.ed.)
  - $Sch6_i$ – vector of characteristics of school attended by student i, (Catholic/public, school size)
  - $Census_i$ – set of census characteristics of neighbourhood of student i

Key variables

- S_i – 1 if "opted out" of assigned high school
  - Manually determined assigned high school for each elementary school in sample
  - In some cases in TDSB, there are multiple assigned schools – if student attended any of them
  - In TCDSB, there is no assigned school
    - Used two definitions: nearest high school, and most attended high school
- Choice measures - Counts of accessible schools
  - A high school is accessible to the student if his/her residence falls within the (empirical) school travel zone
  - Use either total count of accessible high schools, count of public high schools, or count of catholic high schools

School Travel Zones

- Based on actual attendance data of school following methodology of Gibbons, Machin and Silva (2008) – endogenous measure of school choice
- Area around school is divided into 10 “pie-shaped” wedges, each containing 10% of school’s student body
- Each wedge ends at distance of 75th percentile student within that wedge

Travel zones (mostly) do not cross school board boundaries

- Students select from school nearby their residence within their board
  - Less than 1% cross board boundaries
- Travel zone boundaries are at 75th percentile of student distances
  - So, in a given direction, 25% of students would have to be beyond the board boundary for the travel zone to cross

Travel zone example – Lincoln Alexander School (Peel-Toronto Boundary)

Average High School Travel Zone Distances

	Number of schools	Percentage
Less than 2 km	50	13.2%
2 to 4 km	183	48.2%
4 to 6 km	73	19.2%
6 to 8 km	34	8.9%
8 to 12 km	21	5.5%
12 to 15 km	10	2.6%
15 or more km	9	2.4%
Total	380	100.0%
Mean km, (St. Dev)	4.6	(3.40)
Min km, Max km	0.6	24.0

IV methodology – Distance from School Board Boundary

- Should be concerned about endogeneity of accessible school count variables
- Use distance from school board boundary as instrument for accessible school count
  - Since school travel zones don't cross board boundaries, the closer one lives to a board boundary, the fewer accessible travel zones (easily passes F test)
  - Little reason to think that distance from board boundary otherwise impacts likelihood of opting out

Table 1 - Number of students opting out of their Assigned High Schools

School Board (at grade 6)	Total Students	Went to Other Than Assigned HS	Switch denominations
Dufferin Peel Catholic DSB	10,616	25.3%	6.4%
Peel District School	16,726	43.1%	14.2%
Durham Catholic DSB	3,497	22.2%	10.7%
Durham DSB	8,300	22.7%	5.0%
Toronto Catholic DSB (nearest school)	11,410	71.4%	7.5%
Toronto DSB (include tech/comm)	29,728	46.0%	5.6%
York Catholic DSB	6,778	22.1%	9.6%
York Region DSB	12,803	29.7%	6.3%
Total	99,858	41.0%	7.8%

Table 2 - Grade 6 Test Scores by School Board			
	Number of Students / Number of Elementary Schools	Gr 6 Math Scores	Gr 6 Reading Scores
Dufferin Peel Catholic DSB	10,616	2.70	2.77
	112		
Peel District School	16,726	2.73	2.72
	84		
Durham Catholic DSB	3,497	2.71	2.75
	41		
Durham DSB	8,300	2.70	2.74
	95		
Toronto Catholic DSB	11,410	2.66	2.66
	167		
Toronto DSB	29,728	2.77	2.73
	313		
York Catholic DSB	6,778	2.91	2.90
	75		
York Region DSB	12,803	2.96	2.86
	126		
Total	99,858	2.77	2.75
	1013	(0.75)	(0.70)

Table 2 (cont.) - Average Number of Accessible Schools by Board			
	Total Accessible High Schools	Accessible Public High Schools	Accessible Catholic High Schools
Dufferin Peel Catholic DSB	8.67	4.75	3.92
	(3.64)		
Peel District School	8.46	4.83	3.63
	(3.68)		
Durham Catholic DSB	3.58	2.30	1.28
	(1.54)		
Durham DSB	4.02	2.78	1.24
	(1.59)		
Toronto Catholic DSB	18.91	12.25	6.66
	(7.93)		
Toronto DSB	18.98	12.66	6.33
	(8.03)		
York Catholic DSB	4.38	2.35	2.02
	(1.61)		
York Region DSB	5.01	2.85	2.16
	(1.78)		
Total	11.55	7.32	4.23
	(8.48)	(6.19)	(2.77)

Linear Probability Regressions for Likelihood of Opting out of Assigned School				
	(1)	(2)	(3)	(4)
VARIABLES	Math Scores	Math scores – with Gr 6 school FE	Reading Scores	Reading scores – with Gr 6 school FE
Gr. 6 Math Score	0.018***	0.025***	--	--
	(0.004)	(0.003)	--	--
Gr. 6 Reading Score	--	--	0.020***	0.026***
	--	--	(0.004)	(0.003)

Other significant control variables

- Positive and strongly significant – female, special ed., gifted, ESL, french immersion, population density, TCDSB dummy
- Negative and significant control variables: Average household income, Other Board dummies (TDSB omitted)
- Also ran regressions within boards – strongest effect of ability in boards with most choice

Linear Probability Regressions for Likelihood of Opting Out of Assigned School				
Including School Choice Variables				
	(1)	(2)	(3)	(4)
VARIABLES	Math scores	Math scores with interaction	Reading scores	Reading scores with interaction
Accessible High School Count	0.001	0.002	0.001	-0.002
	(0.001)	(0.002)	(0.001)	(0.002)
Gr. 6 Math Score	0.018***	0.005	--	--
	(0.004)	(0.006)	--	--
Gr. 6 Reading Score	--	--	0.020***	0.005
	--	--	(0.004)	(0.006)
School Count * Test Score	--	0.001**	--	0.001***
	--	(0.000)	--	(0.000)

Note: Robust standard errors are clustered on grade 6 school.

Table 7B Summary Linear probability with IV Regressions Using				
Distance to Board Boundary as Instrument – Math Scores				
	(1)	(2)	(3)	(4)
VARIABLES	Instrument for School Choice	Instrument for School Choice – add ability	Instrument for School Choice – add choice*ability	Instrument for Choice * Ability
Accessible High School Count	0.013***	0.013***	0.058***	-0.312***
	(0.001)	(0.001)	(0.007)	(0.055)
Gr. 6 Math Score	--	0.020***	0.189***	-1.188***
	--	(0.004)	(0.023)	(0.209)
School Count * Test Score	--	--	-0.016***	0.117***
	--	--	(0.002)	(0.020)

Table 8B Summary Linear Probability Regressions Using Distance to Board Boundary as Instrument – Reading Scores			
	(1)	(2)	(3)
VARIABLES	Instrument for School Choice – add ability	Instrument for School Choice – add choice*ability	Instrument for Choice * Ability
Accessible High School Count	0.013***	0.060***	-0.417***
	(0.001)	(0.007)	(0.086)
Gr. 6 Reading Score	0.023***	0.203***	-1.614***
	(0.004)	(0.024)	(0.332)
School Count * Test Score	--	-0.0175***	0.158***
	--	(0.002)	(0.032)

Summary of Results

- Strong evidence that more able students (based on test scores) are more likely to opt out of their assigned high school
- Strong evidence that more choice leads to more opting out (not surprising)
- Some evidence that more choice impacts more able students more strongly than students of lesser ability
- Some evidence that students opting out head to schools with stronger peer groups

Discussion

- It seems that increased school choice leads to greater sorting by ability
  - If peer effects are important, this could be negative for students left at weaker schools
- Remain agnostic on peer effects, so cannot say that this is necessarily a bad thing
  - Schools could specialize in teaching students of a given ability

Future work on school choice

- Expand analysis outside of GTA
- Expand analysis of school impacts
- Study longer-run impacts on students (e.g. – outcomes at university)